Contents

1 Introduction: The domain 1

1.1 Aims 1

1.2 Definition 13

1.2.1 Semantic properties 16

1.2.2 Functional properties 18

1.2.3 Form-related properties 23

1.3 Corpora 27

1.3.1 The German corpora 29

1.3.2 The English corpora 34

1.4 Methods 35

1.4.1 Interpretative methods 36

1.4.2 Quantitative and computational methods . . 55

1.4.3 Linguistic models 58

1.5 The structure of the following 60

2 Contexts and categories: Functional interpretation 65

2.1 The functional spectrum of German *ja* 68

2.2 Category assignment 97

2.2.1 The descriptive inventory 99

2.2.2 Classification in artificial neural networks . 110
2.3 Consequences for lexical representation: Constructions 120

3 Conceptual background frame: Evidence from extra-linguistic variables 125
3.1 The variable *communication partner* 126
3.2 The variable *speaker's gender* 157
3.2.1 *Äh* and *ähm* in human-to-human communication 157
3.2.2 Gender-related functional shifts in human-computer interaction 166
3.3 Consequences for lexical representation: Conceptual background frame 175
3.3.1 The relation lexeme – function 176
3.3.2 A frame of communicative domains 178

4 Lexical analysis 195
4.1 Semantic relations 197
4.1.1 Translation equivalents 199
4.1.2 Semantic fields 207
4.2 Semantic decomposition 219
4.2.1 Methodological considerations 219
4.2.2 Semantic tests for discourse particles . . . 223
4.2.3 English *oh* 226
4.2.4 Tests for the features of *oh* 232
4.2.5 Further English discourse particles 239
4.3 Consequences for lexical representation: Invariant meanings 258

5 Lexical representation 261

5.1 A unified model of the meanings and functions of discourse particles 261

5.1.1 The contextual meanings of discourse particles 262

5.1.2 From contextual meanings to discourse functions 270

5.1.3 The different word classes 275

5.1.4 The general function of discourse particles 277

5.2 Aspects of the lexicon 284

5.2.1 General properties of linguistic lexica . . . 285

5.2.2 The structure of lexical entries 289

5.2.3 Types of lexical information 292

5.2.4 Linguistic generalisations in ILEX/DATR . 296

5.3 A frame- and construction-based lexicon for discourse particles 300

6 Conclusion and prospects 321

6.1 From cognitive semantics to lexical pragmatics . . 321

6.2 Automatic processing of discourse particles 323

References 327

Appendix A: Questionnaire 357

Appendix B: DATR Program 359

Index 367